XPACE TECHNOLOGIES

Task Title: Customer Segmentation and Sales Analysis

Scenario:

XPACE TECHNOLOGIES Pvt Ltd has been approached by a retail company that wants to better understand its customer base and improve its sales strategy. They have provided you with a dataset containing their sales data and customer information. Your task is to analyze the data, segment the customers into different groups based on their purchasing behavior, and provide actionable insights that can help the company increase its sales.

Objective:

1. Segment customers into different groups based on their purchasing patterns.
2. Analyze the sales data to identify trends and patterns.
3. Create visualizations to represent your findings.
4. Provide a report with recommendations for the company.

Dataset:

You are provided with a CSV file containing the following columns:

- **Customer ID**: Unique identifier for each customer.
- **Order ID**: Unique identifier for each order.
- **Order Date**: Date when the order was placed.
- **Product ID**: Unique identifier for each product.
- **Product Name**: Name of the product.
- **Category**: Category of the product.
- **Quantity**: Number of units sold.
- **Unit Price**: Price per unit of the product.
- **Total Amount**: Total amount for the order (Quantity x Unit Price).
- **Region**: The region from where the order was placed.

Tasks:

1. **Data Cleaning**:
    - Handle any missing values in the dataset.
    - Standardize the date format for consistency.
    - Remove any duplicate records to ensure data accuracy.
2. **Customer Segmentation**:
    - Use RFM (Recency, Frequency, Monetary) analysis to segment customers into different groups based on their purchasing behavior.
    - Identify key customer segments such as "High-Value Customers," "At-Risk Customers," and "New Customers."
3. **Sales Trend Analysis**:
    - Analyze monthly and yearly sales trends.
    - Determine which product categories are performing best.
    - Identify the regions with the highest sales.
4. **Data Visualization**:
    - Create visualizations using Matplotlib or Seaborn to represent customer segments and sales trends.
    - Include a line chart for sales over time, a bar chart for top product categories, and a pie chart for regional sales distribution.
5. **Recommendation Report**:
    - Write a report summarizing your findings.
    - Provide actionable recommendations for the retail company based on your analysis.
    - Suggest strategies for targeting different customer segments to increase sales.

Tools Required:

- Python (Pandas for data manipulation, Matplotlib and Seaborn for visualizations)
- Jupyter Notebook or Google Colab for coding and analysis

Expected Outcome:

- Cleaned and segmented customer data.
- Clear visualizations representing key insights.
- A well-written report with findings and recommendations.

Submission Guidelines:

- Submit your code, dataset, and report via a GitHub repository.
- Ensure that your notebook is well-documented with comments explaining each step.
- The deadline for this task is **10-September-2024**.

Evaluation Criteria:

- **Data Cleaning**: Accuracy in handling missing values and duplicates.
- **Segmentation and Analysis**: Effectiveness in customer segmentation and identification of sales trends.
- **Visualizations**: Clarity and relevance of the visualizations created.
- **Report**: Quality of insights and actionable recommendations provided.